

Serial No. 10/006,492

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

1 - 142. (Canceled)

143. (New) A method for determining a probability that a set of biopolymer sequences is a member of a hidden Markov model (HMM) of a set of known biopolymer sequences comprising the steps of:

a) constructing said HMM by

- 1) identifying at least two regions in the set of known biopolymer sequences and determining a state of matching for each region using scoring matrices, wherein at least one state of matching is similarity and at least one state of matching is dissimilarity;
- 2) producing at least one module for each of said match states from one or more nodes linked in an order, wherein each node represents a distribution of monomers at corresponding positions in the sequences of said set of known biopolymer sequences;
- 3) linking said modules in the order in which regions occur to form the HMM; and

b) comparing a second set of biopolymer sequences to said HMM by evaluating the scoring matrices for the second set of biopolymer sequences to produce an output score that is the probability that the second set of biopolymer sequences is a member of the HMM.

144. (New) The method of claim 143, wherein said determining said state of matching comprises determining said state of matching against a reference sequence.

145. (New) The method of claim 143, wherein one of said scoring matrices is an arithmetic inverse function of another one of said scoring matrices.

Serial No. 10/008,492

146. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises sequences from different species.

147. (New) The method of claim 146, wherein said different species comprise mammalian species.

148. (New) The method of claim 143, wherein said set of known biopolymer sequences consists of two sequences.

149. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises at least three sequences.

150. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises amino acid sequences.

151. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises nucleic acid sequences.

152. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises genomic nucleic acid sequences.

153. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises a plurality of non-coding regions.

154. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises a plurality of regulatory regions.

155. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises a plurality of transcriptional regulatory regions.

Serial No. 10/005,492

156. (New) The method of claim 143, wherein said distribution of monomers at a first node is different from said distribution of monomers at a second node.

157. (New) The method of claim 143, wherein one of said scoring matrices is a function of independent probabilities of a monomer occurrence.

158. (New) The method of claim 143, wherein said set of known biopolymer sequences comprises prohormones.